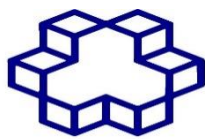


به نام او



دانشگاه صنعتی خواجه نصیرالدین طوسی
گروه مهندسی برق

عنوان :

امتحان پایانترم یادگیری ماشین

دانشجو:

آرمان مرزبان

۴۰۲۰۰۹۱۶

استاد:

دکتر علیاری

ماه و سال

تیرماه ۱۴۰۳

فهرست مطالب

عنوان	صفحه
بخش اول.....	۱
۱-۱- سوال هماهنگی اول.....	۲
۱-۲- سوال هماهنگی دوم.....	۳
۱-۳- سوال ج.....	۵
با استفاده از کد زیر مقدار ضرایب بهینخ و بردادهای پشتیبان بدست می آیند:.....	۷
فصل ۲.....	۸
۲-۱- سوال دوم.....	۹
فصل ۳.....	۱۳
۳-۱- سوال سوم.....	۱۴

فهرست شکل ها

- شکل ۱: نمودار هیستوگرام داده های آموزشی ۹
- شکل ۲: نمودار هیستوگرام داده های آزمون ۱۰
- شکل ۳: ماتریس همبستگی کل ویژگی ها ۱۱
- شکل ۴: ماتریس درهمریختگی بر روی داده های تست با SVM ۱۲

بخش اول

سوال هماهنگی و تحلیلی

۱-۱- سوال هماهنگی اول

اثبات کنید که اگر به جای $CF(\sum_{i=1}^t \xi_i^0)^k$ قرار دهیم $\sum_{i=1}^t \xi_i^2$ ، مقدار بهینه C عبارت است از $0 \leq \alpha_i \leq C$

$$\frac{\partial L}{\partial \mathbf{w}} \Big|_{\mathbf{w}=\mathbf{w}_0} = \mathbf{w}_0 - \sum_{i=1}^t \alpha_i y_i \mathbf{x}_i = 0,$$

$$\frac{\partial L}{\partial b} \Big|_{b=b_0} = \sum_{i=1}^t \alpha_i y_i = 0,$$

$$\frac{\partial L}{\partial \xi_i} \Big|_{\xi_i=\xi_i^0} = kC \left(\sum_{i=1}^t \xi_i^0 \right)^{k-1} - \alpha_i - r_i.$$

اگر داشته باشیم:

$$\sum_{i=1}^t \xi_i^0 = \left(\frac{\delta}{Ck} \right)^{\frac{1}{k-1}},$$

می توانیم معادله سوم را به صورت زیر بازنویسی کنیم

$$\delta - \alpha_i - r_i = 0.$$

بنابراین خواهیم داشت:

$$\begin{aligned} \mathbf{w}_0 &= \sum_{i=1}^t \alpha_i y_i \mathbf{x}_i, \\ \sum_{i=1}^t \alpha_i y_i &= 0, \\ \delta &= \alpha_i + r_i. \end{aligned}$$

با جایگذاری \mathbf{w}_0, b_0 و δ داریم:

$$W(\Lambda, \delta) = \sum_{i=1}^t \alpha_i - \frac{1}{2} \sum_{i=1}^t \sum_{j=1}^t \alpha_i \alpha_j y_i y_j \mathbf{x}_i \cdot \mathbf{x}_j - \frac{\delta^{k/k-1}}{(kC)^{1/k-1}} \left(1 - \frac{1}{k}\right)$$

برای یافتن راهحل ابرصفحه حاشیه نرم، باید فرم تابعی (۵۹) را تحت محدودیت‌های (۵۷)-(۵۸) با توجه به متغیرهای غیر منفی α_i, r_i با $i = 1, \dots, t$ به حداکثر برسانیم. نماد برداری (۵۹) را می‌توان به صورت بازنویسی کرد

$$W(\Lambda, \delta) = \Lambda^T \mathbf{1} - \left[\frac{1}{2} \Lambda^T \mathbf{D} \Lambda + \frac{\delta^{k/k-1}}{(kC)^{1/k-1}} \left(1 - \frac{1}{k}\right) \right],$$

جایی که Λ و \mathbf{D} همانطور که در بالا تعریف شده است. بنابراین برای یافتن نقطه زینی مورد نظر باید ماکزیمم رابطه قبل را تحت محدودیت‌ها پیدا کرد

$$\begin{aligned} \Lambda^T \mathbf{Y} &= 0 \\ \Lambda + \mathbf{R} &= \delta \mathbf{1} \\ \Lambda &\geq 0 \end{aligned}$$

9

$$\mathbf{R} \geq 0.$$

بنابراین مقدار Λ که روابط فوق را برقرار می‌کند به صورت زیر است:

$$0 \leq \Lambda \leq \delta \mathbf{1}.$$

بنابراین با توجه به رابطه $\sum_{i=1}^t \xi_i^0 = \left(\frac{\delta}{Ck}\right)^{\frac{1}{k-1}}$ خواهیم داشت:

$$\delta = Ck \left(\sum_{i=1}^t \xi_i^0 \right)^{k-1}$$

به ازای $k = 1$ رابطه اخیر به صورت زیر بدست می‌آید و اثبات کامل می‌گردد

$$0 \leq \Lambda \leq C.$$

۱-۲- سوال هماهنگی دوم

سوال ۲ هماهنگی: اثبات کنید که $S_T = S_w + S_B$

حل: با بازنویسی روابط قبلی می‌توانیم بنویسیم:

سؤال 2: همبستگی: اثبات کنید که $S_T = S_W + S_B$
 حل: با باز نویسی روابط همبستگی که تعریف شده داریم:

$$1 \quad S_W = \sum_{c \in \text{class}} \sum_{x_j \in c} (x_j - \mu_c)(x_j - \mu_c)^T$$

$$2 \quad S_B = \sum_{c \in \text{class}} n_c (\mu_c - \mu)(\mu_c - \mu)^T$$

$$3 \quad S_T = \sum_x (x - \mu)(x - \mu)^T$$

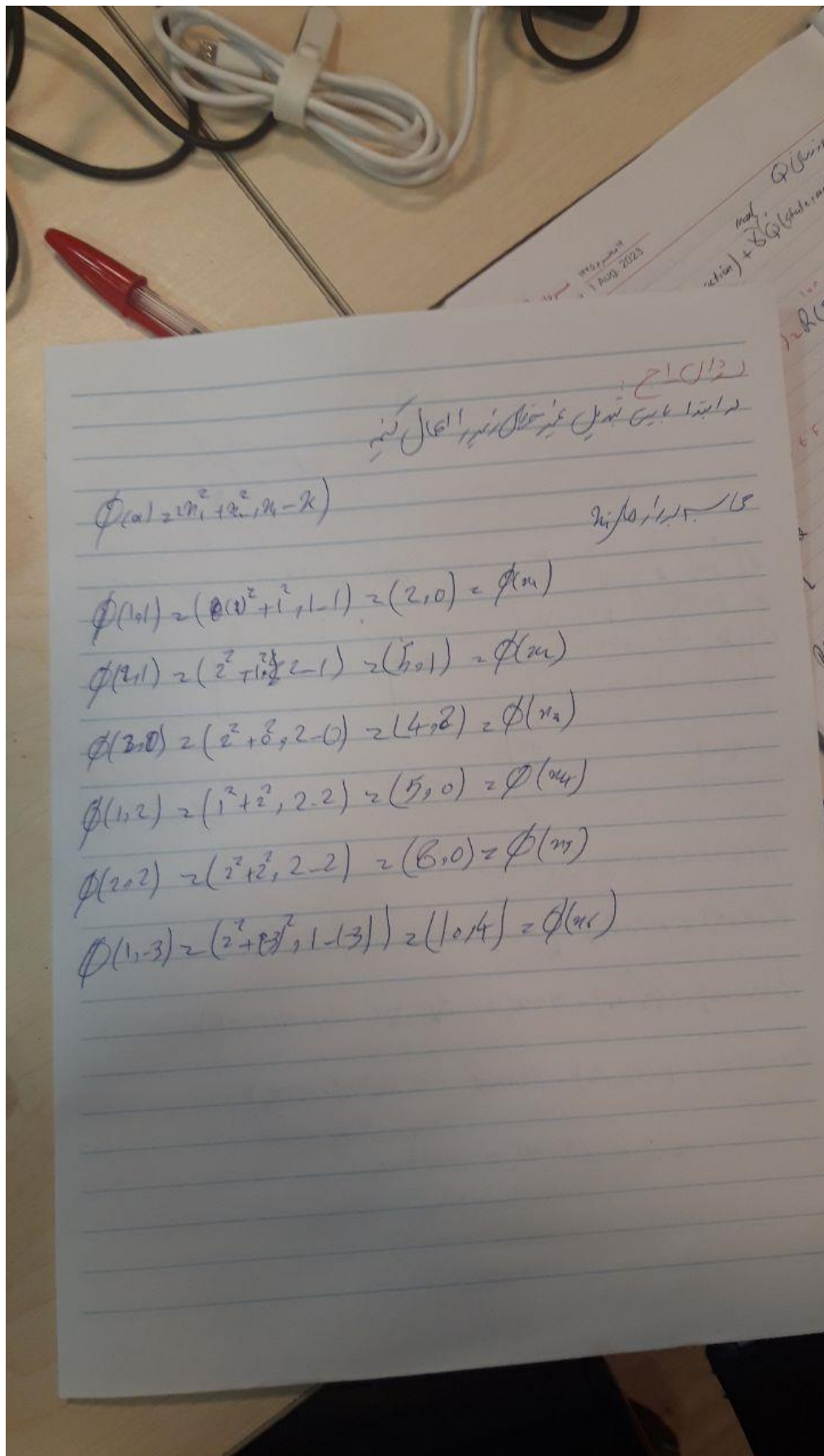
با باز نویسی رابطه اول سؤال داریم:

$$S_T = S_W + S_B$$

$$S_T = \sum_{c \in \text{class}} \sum_{x_j \in c} (x - \mu_c + \mu_c - \mu)(x - \mu_c + \mu_c - \mu)^T = \sum_{c \in \text{class}} \sum_{x_j \in c} (x - \mu_c)(x - \mu_c)^T + \sum_{c \in \text{class}} \sum_{x_j \in c} (\mu_c - \mu)(\mu_c - \mu)^T$$

$$(x - \mu_c)^T = S_W + \sum_{c \in \text{class}} n_c (\mu_c - \mu)(\mu_c - \mu)^T = S_W + S_B$$

۱-۳- سوال ج



با استفاده از کد زیر مقدار ضرایب بهینه و بردادهای پشتیبان بدست می آیند:

```
import numpy as np
from sklearn import svm

#
X = np.array([[2, 0], [5, 1], [4, 2], [5, -1], [8, 0], [10, 4]])
y = np.array([1, 1, 1, -1, -1, -1])

#
clf = svm.SVC(kernel='linear', C=1.0)
clf.fit(X, y)

#
support_vectors = clf.support_vectors_
coefficients = clf.coef_
intercept = clf.intercept_

print("Support Vectors:")
print(support_vectors)
print("Coefficients:")
print(coefficients)
print("Intercept:")
print(intercept)
```

Support Vectors:

[[5. -1.]

[10. 4.]

[5. 1.]]

Coefficients:

[[-0.99975716 0.99959527]]

Intercept:

[4.99892073]

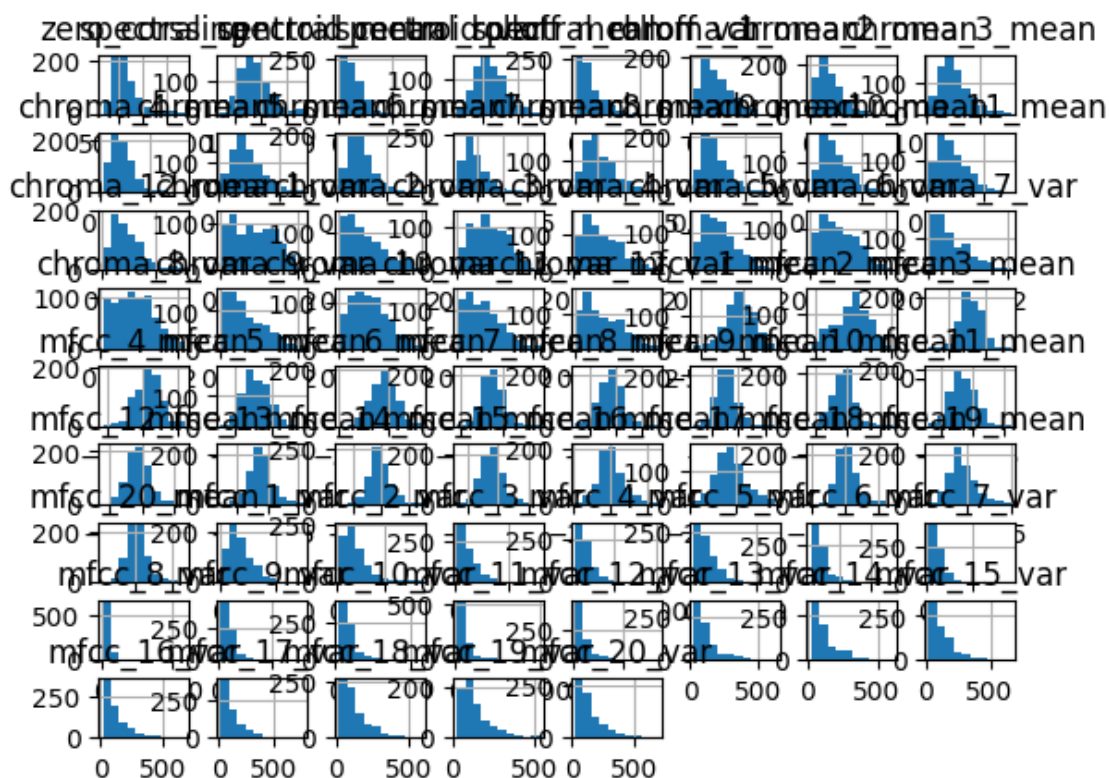
فصل ۲

سوال دوم

۲-۱- سوال دوم

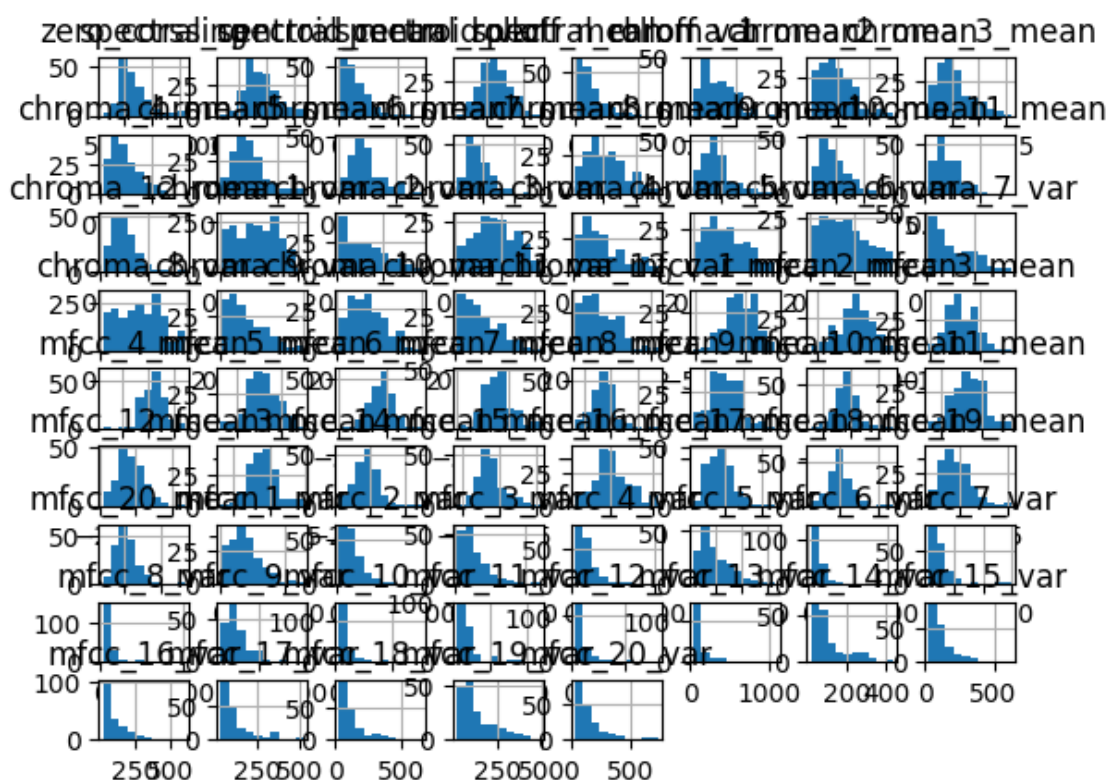
(آ)

نمودار هیستوگرام داده های آموزشی به صورت شکل زیر است:



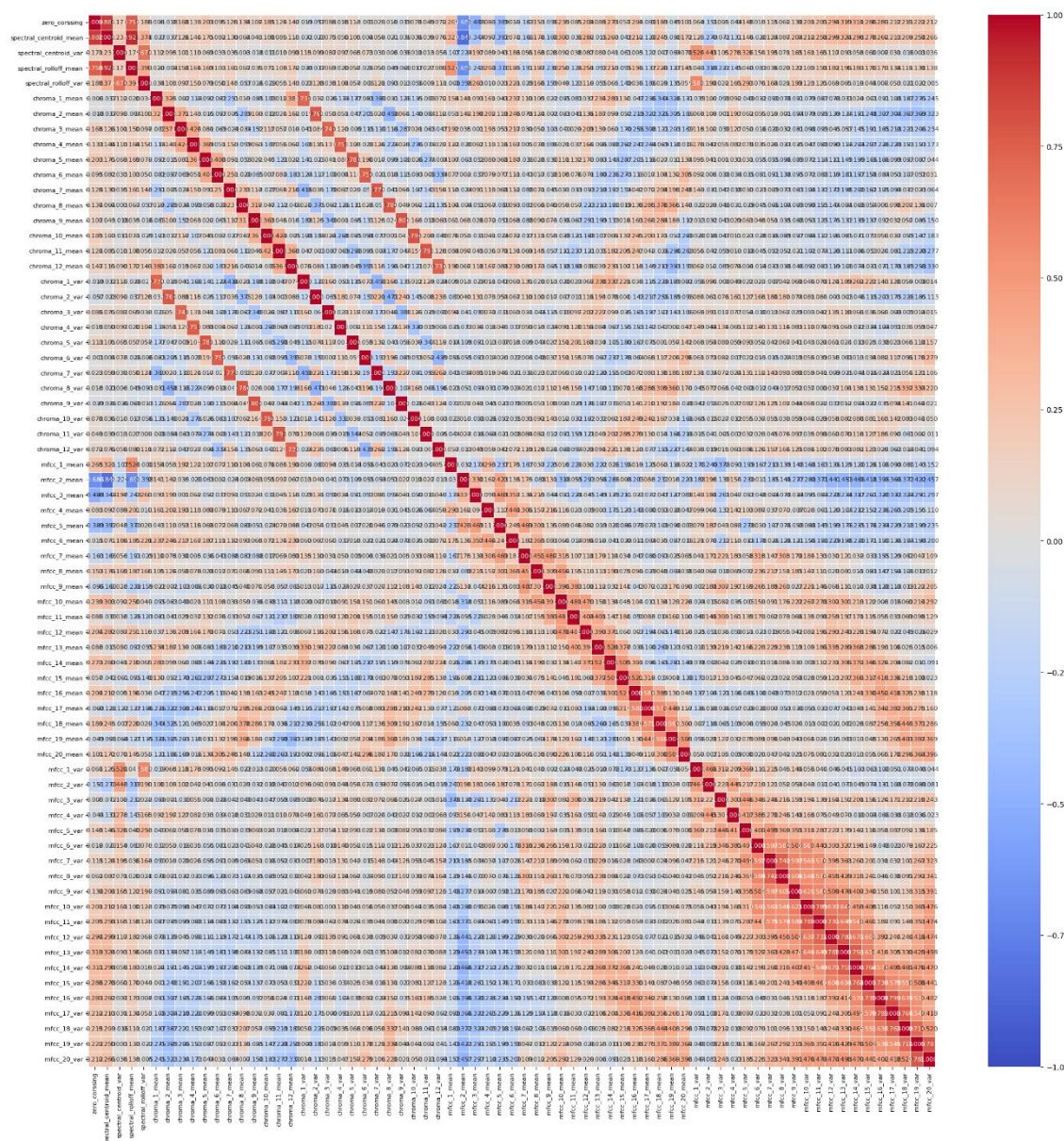
شکل ۱: نمودار هیستوگرام داده های آموزشی

نمودار هیستوگرام داده های تست یا آزمون نیز به صورت زیر است:



شکل ۲: نمودار هیستوگرام داده های آزمون

همانطور که از شکل ۱ و ۲ پیداست توزیع داده ها به صورت نرمال هم در بخش آزمون و هم ارزیابی است. ماتریس همبستگی نیز به صورت شکل زیر است:



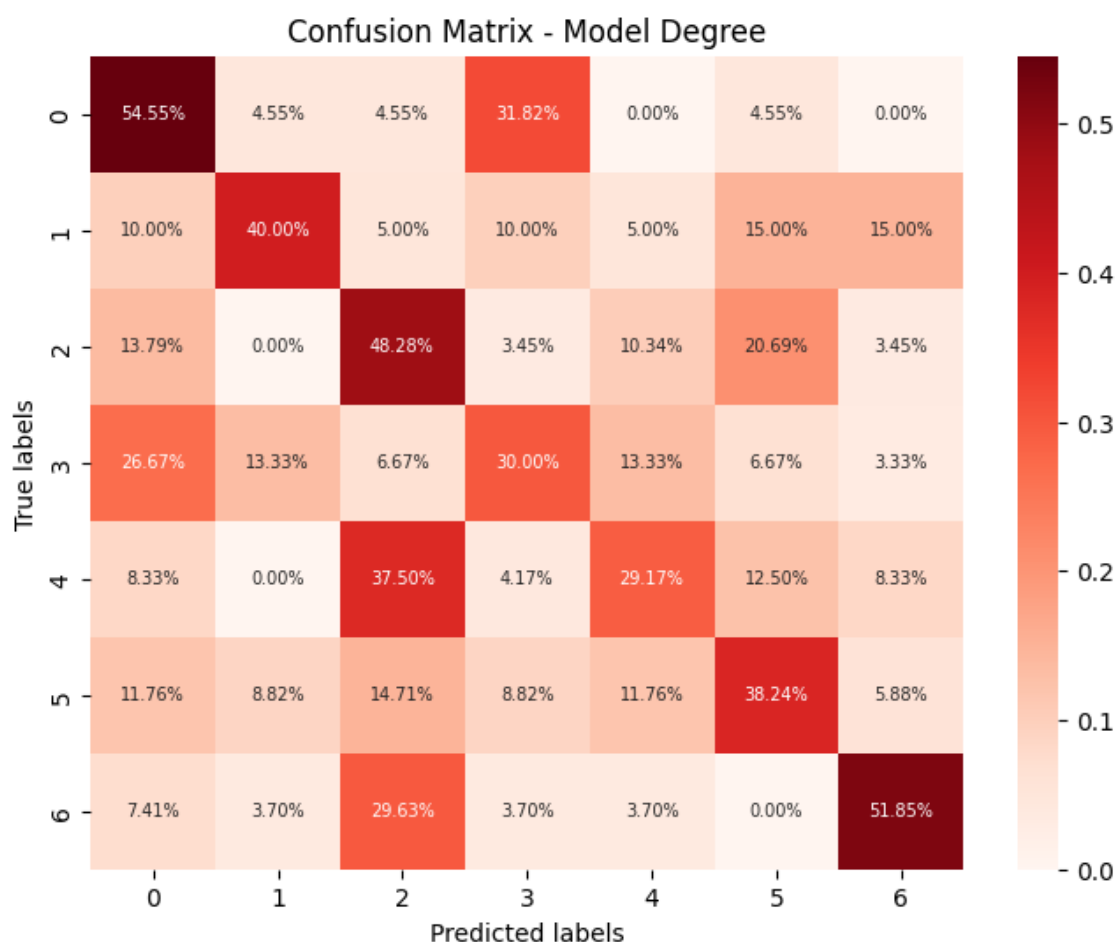
شکل ۳: ماتریس همبستگی کل ویژگی ها

همانطور که مشاهده می شود در مواردی که میزان همبستگی زیاد است رنگ خانه ها بسیار قرمز تر است و از طرفی چون تعداد ویژگی ها ۷۲ است به راحتی از روی شکل نمی توان تحلیل درستی انجام داد اما بصورت شهودی واضح است که بسیاری از ویژگی ها به هم همبسته هستند و بایستی کاهش بعد صورت بگیرد.

با استفاده از قطعه کد زیر بعد داده ها را به مقدار ۲۰ کاهش داده ایم:

```
from sklearn.datasets import load_breast_cancer
from sklearn.feature_selection import GenericUnivariateSelect,
chi2
transformer = GenericUnivariateSelect(chi2, mode='k_best',
param=20)
X_new = transformer.fit_transform(X, y)
X_new.shape
```

با استفاده از کاهش بعد به روش PCA بر روی داده ها به میزان ۳۰ نتیجه آموزش مدل SVM با کرنل rbf و دقت ۴۱ درصدی بر روی داده های تست ماتریس در همریختگی به این صورت حاصل شده است:



شکل ۴: ماتریس درهمریختگی بر روی داده های تست با SVM

در این حالت مطابق شکل ۴ مشاهده می شود که مدل SVM به خوبی آموزش ندیده است و دقت آن بسیار ضعیف است لذا با بررسی پارامترهای مختلف مانند بعد، و هایپر پارامترهای دیگر می توان عملکرد آن را بهبود داد.

لینک کد این بخش در این [مسیر](#) است.

فصل ۳

سوال سوم

۳-۱- سوال سوم

شایان ذکر است که تمامی المان های ماتریس Q را می توان با استفاده از قانون مارکوف آپدیت کنیم.